CP 23871 (5)

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : H04J 3/16	A1	(11) International Publication Number: WO 99/22477 (43) International Publication Date: 6 May 1999 (06.05.99)

(21) International Application Number: PCT/US97/19207

(22) International Filing Date: 24 October 1997 (24.10.97)

(71) Applicant (*for all designated States except US*): THE TRUSTEES OF COLUMBIA UNIVERSITY IN THE CITY OF NEW YORK [US/US]; Broadway & 116th Street, New York, NY 10027–6699 (US).

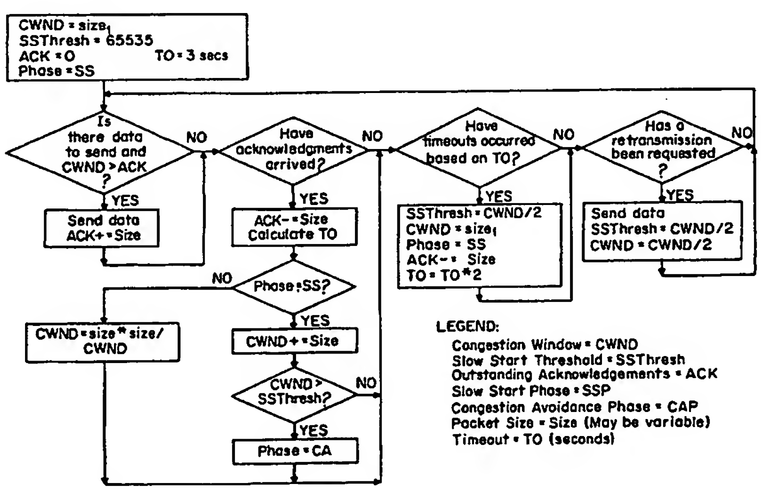
(72) Inventors; and
(75) Inventors/Applicants (*for US only*): JACOBS, Stephen [US/US]; Columbia University, Dept. of Electrical Engineering, 530 West 120th Street, New York, NY 10027–6699 (US). ELEFTHERIADIS, Alexandros [GR/US]; Columbia University, Dept. of Electrical Engineering, 530 West 120th Street, New York, NY 10027–6699 (US).

(74) Agents: TANG, Henry et al.; Brumbaugh, Graves, Donohue & Raymond, 30 Rockefeller Plaza, New York, NY 10112–0228 (US).

(81) Designated States: CA, US.

Published
With international search report.

(54) Title: TRANSMISSION CONTROL FOR MINIMIZING CONGESTION IN DIGITAL COMMUNICATIONS NETWORKS



(57) Abstract

In congestion control in a digital communications network such as the Internet or corporate "Intranets", and especially in real–tine transmissions in such networks, perfect reliability may not be required. For increased likelihood that data arrive on time, an estimate is used of the bandwidth which is available from a sender to a receiver. The estimate is increased or decreased, by the sender, depending on monitoring of acknowledgemnts from the receiver. The technique coexists well with protocols based on TCP (Transmission Control Protocol), such as FTP (File Transfer Protocol) and HTTP (Hyper Text Transfer Protocol), by sharing the available bandwidth equally.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

TRANSMISSION CONTROL FOR MINIMIZING CONGESTION IN DIGITAL COMMUNICATIONS NETWORKS

Technical Field

The invention relates to transmissions in a digital communications network and, more specifically, to transmission control for minimizing network congestion.

Background of the Invention

For preventing loss of data due to congestion in digital network communications, a protocol known as Transmission Control Protocol (TCP) has been proposed for the Internet; see Information Sciences Institute, "Transmission Control Protocol - Request for Comments 793", September 1981 and W. Stevens, "TCP Slow Start, Congestion Avoidance, Fast Retransmit, and Fast Recovery Algorithms - Request for Comments 2001", January 1997. TCP is based on the notion of fair sharing of transmission resources among users.

TCP is reliable, in the sense that the data received at a destination are an exact duplicate of the data that was sent. Such reliability may be at the expense of transmission delays, however.

For some transmissions, e.g. real-time audio and video, reliability is less important, and the primary concern is with the data arriving on time. Specifically, for example, it is more acceptable to lose an occasional frame of video than to have the video start and stop repeatedly.

Summary of the Invention

For congestion control in a digital communications network such as the Internet or corporate "Intranets", and especially for real-time transmissions in such networks,

a transmission technique is preferred which is not perfectly reliable, but which makes it more likely that the data arrive on time. The technique uses an estimate of the bandwidth which is available in a network, from a sender to a receiver. The estimate is increased or decreased, by the sender, depending on monitoring of acknowledgments from the receiver.

The technique is compatible with TCP, and its use by a sender in a connection results in fair sharing of network resources with all other connections. It can be used, e.g., with well-established protocols such as File Transfer Protocol (FTP) and Hyper Text Transfer Protocol (HTTP).

Brief Description of the Drawing

Fig. 1 is a representation of packet format for a preferred embodiment of the invention.

Fig. 2 is a flow chart for processing at a network server, in accordance with a preferred embodiment of the invention.

Figs. 3a and 3b are schematics of communications systems in accordance with preferred embodiments of the invention, with fixed and adaptable bandwidth requirements, respectively.

Fig. 4 is a flow chart for exemplary rate control processing in a system according to Fig. 3b.

Fig. 5a is a graphic representation of system behavior for an example in a system in accordance with Fig. 3a.

Fig. 5b is a graphic representation of system behavior for an example in a system in accordance with Fig. 3b.

Fig. 6 is a representation of packet format for a preferred embodiment of the invention in a wireless or hybrid wired-wireless network.

Detailed Description of Preferred Embodiments

While preferred embodiments are described in the following primarily in method terms, the inventive technique also includes systems embodiments, e.g. involving a programmed processor. A prototype implementation uses a Unix Workstation as network server and a PC as client server, both programmed in C++. Use of special-purpose firmware or hardware is not precluded.

The technique is window-based in the sense that a sender maintains a count of the number of outstanding packets, i.e., packets which have been sent, but for which an acknowledgment has not yet been received from the receiver. The sender maintains current an upper bound on the number of outstanding packets allowed in the network, called the "congestion window" (CWND) and providing an indication of the available bandwidth from sender to receiver. Congestion is detected when a packet is lost in the network. Alternatively, and especially in transmissions of variable-length packets, CWND can be maintained in units of bytes rather than units of packets.

If the number of outstanding packets is less than CWND, the sender can continue to send data into the network. Otherwise, the sender must stop transmitting data until either CWND increases or the number of outstanding packets decreases. If acknowledgments are received, CWND will increase, and the number of outstanding packets will decrease. If no acknowledgments are returned, packets will timeout and be deemed lost by

the protocol, thus decreasing the number of outstanding packets.

Optionally, selective retransmission can be provided for. A current estimate is maintained of the round trip time, i.e. the time elapsed between sending a packet and receiving an acknowledgment. The protocol sends the estimate to the receiver in each packet header. When the receiver determines that a packet has been lost, it then determines if there is enough time to receive the retransmitted packet before it is needed. If so, the receiver can request a retransmission; otherwise, no request is made. In real-time audio or video, for example, if the receiver has 100 milliseconds worth of data buffered for playback when detecting loss of a packet, and if the estimate for the round-trip time is less than 100 milliseconds, a request for retransmission is likely to result in timely retransmission of the lost packet. Thus, a best-effort attempt is made at reliability.

As illustrated by Fig. 1, a data packet includes the standard User Datagram Protocol (UDP) header, a 2-byte sequence number, a 4-byte time stamp, and a 4-byte round-trip time estimate measured in milliseconds. The sequence number is for packet reordering at the receiver, in case packets arrive out of order. The time stamp is media dependent and generally provides an indication of the presentation time of the packet.

Fig. 2 illustrates preferred packet processing by a server system. There is a main loop which continually checks whether (i) data can be sent out, (ii) an acknowledgment has arrived, (iii) a timeout has occurred, or (iv) a retransmission was requested. Initially, CWND is set to the size of the first packet to be transmitted, ensuring that the first packet can be sent out.

"Outstanding acknowledgments" (ACK) is set to zero. "Timeout" (TO) is set to 3 seconds, for example, indicating the amount of time not to be exceeded between sending a packet and receiving its acknowledgment. If an acknowledgment is not received in time, the packet is assumed to be lost. The system starts out in a "Slow-Start Phase" indicated by Phase=SS.

Since CWND is the size of the first packet, ACK=0, and there is data available to send (namely the first packet), the first packet is sent into the network. ACK is then increased by the size of the packet sent, representing the number of bytes currently in the network that have not yet been acknowledged. The system then checks whether acknowledgments have arrived. If so, Outstanding Acknowledgments is decreased by the size of the packet to which the acknowledgment refers: ACK = ACK-size. The system then calculates the Round Trip Time (RTT), i.e. the difference between when a packet was sent and when the acknowledgment was received. RTT is used in the calculation of Timeout (TO).

The system maintains an estimate of the round trip time, RTT_{avg}, by using the measured RTT, RTT_i, for each acknowledgment. Following D. Comer, "Internetworking with TCP/IP", 3rd Edition, Simon & Schuster, 1995, pp. 191-230, RTT_{avg} and Timeout (for future use) are calculated as follows:

$$Diff = RTT_i - RTT_{avg}$$
$$RTT_{avg} = RTT_{avg} + Diff/8$$
$$Dev_i = 0.25 \cdot (|Diff| - Dev_i)$$
$$Timeout = RTT + 0.25 + 3 \cdot Dev_i$$

Now, in Slow Start Phase, CWND is increased by *size*:

CWND = CWND + size;

later, in Congestion Avoidance Phase (Phase≠SS), CWND is increased by the square of the size divided by the current value of CWND:

$$CWND = CWND + size^2/CWND.$$

Slow Start calls for increasing the value of CWND each time an acknowledgment is received. In the case of variable length packets, with CWND being the number of bytes of outstanding packets, Slow Start calls for increasing the value of CWND by the size of the packet to which the acknowledgment refers.

After increasing CWND, there follows checking of CWND > SSThresh, the Slow Start Threshold. If true, Phase = CA, for Congestion Avoidance.

Then, concerning timeouts, if an acknowledgment is not received within Timeout (TO) milliseconds after it was sent, the packet is determined to be lost in the network and the appropriate action is taken. This includes (i) setting SSThresh to half of the current CWND, (ii) setting CWND to the value of the next packet to be sent out (i.e. resetting CWND), (iii) setting Phase to Slow Start, (iv) decreasing the outstanding acknowledgment by the size of the packet which timed out, and (v) doubling the Timeout period (TO).

Finally, the system checks for receipt of a retransmission request. If so, it resends the appropriate data and resets SSThresh and CWND to half the current value of CWND. This is known as Fast Recovery. The system then returns to check for further data to send, and whether CWND > ACK.

As described, the technique does not depend on whether the bandwidth requirements of the media can be changed or adapted. Fig. 3a shows a system with non-adaptable media, such as MPEG. The server reads the media from a file or obtains it from a live source and

fills a buffer. At the server, the media pump sends the data to the client from the buffer, taking into account the current value of CWND determined in accordance with Fig. 2, and the media pump supplies the size values for congestion control. In case of significant congestion, CWND will be less than ACK, and this will stop the media pump from sending further data for a period of time, thereby reducing the media pump transmission rate.

So long as the average available bandwidth of a connection is greater than or equal to the bandwidth requirements of the media, and so long as there is sufficient buffering, the media can be played back without interruption. With congestion-minimizing processing as described above, few packets will be lost, and can be retransmitted if there is enough time.

Buffering provides for variation in the available bandwidth: the larger the buffer, the more variation can be accommodated. But there is an initial start-up delay while a client buffer is being filled, so that increased buffering results in a longer start-up delay.

As to adaptable media, there are several ways of changing bandwidth requirements. In the case of MPEG, for example, one way involves dropping frames as described by Z. Chen et al., "Real Time Video and Audio in the World Wide Web", World Wide Web Journal, Vol. 1, January 1996. The server finds the picture header in the MPEG stream and stops sending data until it finds the next picture header in the stream. This has the effect of dropping one frame from the media stream, and thereby reducing the bandwidth requirements. As frames are interdependent in MPEG, a frame should not be dropped if other frames depend on it, i.e. an I-frame cannot be dropped if the stream contains P- or B-frames which depend on it.

For MPEG video, another technique for bandwidth reduction is known as Dynamic Rate Shaping (DRS) as described by A. Eleftheriadis et al., "Constrained and General Dynamic Rate Shaping of Compressed Digital Video", Proceedings, 2nd IEEE International Conference on Image Processing, Washington, D.C., October 1995, pp. III.396-399. This involves identifying, frame by frame, those coefficients in the MPEG stream which are least important in terms of image quality, and removing them from the stream.

Fig. 3b shows an adaptable media system. Again, the original media either is stored locally or is supplied by a live source. But, in this case the data enters a media adaptation module which shapes the media into an estimate of the available bandwidth. The shaped media enters the buffer, which is then read by the media pump. Again, the media pump sends out data so as to comply with the CWND. At the client, the data is buffered for presentation to the user. The client provides feedback information for congestion control.

The status of the buffer between the media adaptation module and the media pump is critical for this system. If the buffer is filling, then the media pump is sending data out more slowly than the media adaptation module is filling the buffer. In this case, the system should decrease the bandwidth requirements of the media so that the buffer does not overflow, by dropping frames or assigning a lower rate to DRS.

Conversely, if the buffer is emptying, the media pump is sending data out faster than the buffer is being filled by the media adaptation module. In this case, the system should increase the bandwidth requirements of the media so that the user gets the best quality possible. Since rate control provides information to the media adaptation module, it is highly dependent on the time of

media being adapted. The media pump operates as in the non-adaptable case, sending data only when CWND > ACK. Based on the occupancy of the buffer, the adaptable media module is instructed to change the rate of the media.

For example, for rate control in MPEG video by frame dropping, a frame can be dropped when the buffer is more than half full; otherwise, the video is passed unaltered to the buffer. Other scenarios, using DRS and more sophisticated rate control may be implemented. For example, if the buffer is filling, the transmission rate may be reduced in inverse relationship to the rate of buffer filling.

Fig. 4 illustrates an exemplary rate control technique based on measurements of buffer occupancy. Every 5 seconds, an average buffer occupancy is obtained for the previous 5 seconds, $Occupancy_i$. The change in the buffer occupancy since the previous 5-second interval, $Occupancy_{i-1}$, is determined as $Diff_i$, Start-up is with $Occupancy_0 = 0$.

The Centering factor provides a weighting for the occupancy to stay close to the desired occupancy at the buffer midpoint. The maximum buffer size is 5 seconds worth of data and depends on the originally encoded rate of the stream.

If $Diff_i < 0$,

$$Centering_i = Occupancy_i/Occupancy_{desired},$$

where $Occupancy_{desired}$ is the buffer occupancy which rate control tries to maintain. Otherwise,

$$Centering_i = 2 - (Occupancy_i/Occupancy_{desired}),$$

the goal being to keep the Centering factor between 0 and 2.

Then, $Beta_i$ is determined as a direct indication of how much demand varies in the network, using the Coefficient of Variation of the past and current values

of the average occupancy. The coefficient of variation is defined as

$$\text{Variance(samples)}/\text{Mean}^2\text{(samples)},$$

where the samples are the two values of the average buffer occupancy. Beta is then multiplied by 10. If Beta is less than 0.1, it is assigned the value 0.1, if it is greater that 1.0, it is assigned the value 1.0.

Finally, the new transmission rate is calculated by subtracting, from the previous rate, the value Beta·Centering·Diff·8, where the factor 8 is due to Diff being in bytes and the rate being in bits. These steps are repeated every 5 seconds.

Adaptable media can cope with more drastic variations in network resources, as compared with non-adaptable media. In non-adaptable media, a decrease in network resources results in less data reaching the receiver than is needed, and the receiver can rely only on its initial buffering to continue playback.

Fig. 5a shows an example of using a non-adaptive media. In this case, the rate of the media is 300 kbps, and the final buffering is 5 seconds (1500 kb). The available bandwidth is continually changing. In the beginning there is just enough bandwidth for the media and no buffering is used. But as soon as the available bandwidth decreases to 200 kbps, the receiver must begin using its buffering. If the bandwidth stays low for an extended period of time, the buffer may become completely depleted, at which time the user will experience an interruption in playback. This occurs at around 40 seconds. The available bandwidth then increases to 350 kbps, at which time the buffer can accumulate again.

With adaptable media, the initial buffering has to be used only when the bandwidth requirements of the media cannot be reduced further. As illustrated by Fig. 5b, for the same rate and initial buffering as in Fig. 5a,

the bandwidth requirements of the media can be reduced down to a minimum of 150 kbps. When the available bandwidth drops to 200 kbps, the media also is reduced to this rate, so that no receiver buffering is used to compensate for the network. However, once the available bandwidth decreases to 100 kbps, the media can only be reduced to 150 kbps, and so the receiver buffer begins to be depleted. This scenario is more robust, as the available bandwidth can drop to 150 kbps and receiver buffering is not used.

Congestion control in accordance with the invention is applicable wherever some degree of loss can be tolerated, including most video and audio codecs, with adaptable codecs being preferred. Most video codecs can be adapted by using frame dropping. Even still images can be adapted for real-time applications. JPEG and MPEG have similarities in the way they are coded, so that a technique like DRS can be used on JPEG as well. A new standard known as Flashpix has the capability to be displayed at different resolutions, and hence different bandwidth requirements when sending a picture across the Internet.

While preferred embodiments have been described above under the assumption of a wired network, composed of fiber-optic or coaxial physical cables, techniques of the invention can be used to advantage with wireless networks as well. As digital communications protocols were originally devised with wired networks in mind, most congestion-aware protocols, TCP included, assume that a lost packet indicates congestion. This is practicable in wired networks, where bit errors are uncommon. Bit errors are more common in a wireless environment, however, so that a packet is more likely to become “lost” due to an error in the packet, regardless of congestion. But known systems do not include facilities for informing

the receiver when a packet has arrived containing an error. Internet Protocol (IP) packets are simply dropped at the receiver if there is an error in the header.

Currently, with UDP, the receiver system has the option of instructing the sender system not to put error checking in packets. This is on a system-wide basis, so that all UDP packets coming from the sender system will not use error checking, which is undesirable when other applications expect UDP error checking.

Preferably, in accordance with a preferred embodiment of the invention, the receiver can distinguish whether a packet is lost due to congestion or error, in an application-specific fashion.

Fig. 6 illustrates a packet constructed from an IP packet provided with the shaded area by the operating system. Error checking will be over the IP header only, so that a bit error there still results in the packet being dropped without notification. However, without error checking over the payload, a bit error in the payload does not result in the packet being dropped.

In this embodiment of the invention, the sender constructs a UDP header inside the payload of the IP packet, for the packet to appear as a regular UDP packet at the receiver. In the UDP header, the sender sets the Cyclic Redundancy Code (CRC) field to zero, indicating that no error checking is used. Accordingly, when the receiver reads the packet, the UDP module of the receiver system will not do any error checking, leaving it to the application to check for errors.

So that packets received with errors are not used, the sender must insert its own error checking functionality into the payload of the UDP packet it constructs. In Fig. 6, this is shown as Application Defined CRC. If, using Application Defined CRC, the receiver determines that there is an error, the receiver

application drops the packet and sends a request for retransmission to the sender— without invoking congestion avoidance to reduce the transmission rate at the sender. If there is no error, the packet is used by the receiver application, with regular acknowledgment.

In this fashion, the likelihood of a packet being dropped by the receiver operating system due to packet error is minimized, and greater throughput is realized on wireless networks without impairing the performance on wired networks. No changes are required to the operating system nor the underlying network link layer, so long as the link layer does not perform error checking over the entire link layer packet.

This preferred technique can be used with all proprietary client-server protocols which are congestion-aware. Such protocols must be proprietary because of changes to both the client and the server. Accordingly, adaptable media applications are preferred.

Claims

1. A method for transmitting data from a sender to a receiver in a digital communications network, comprising:

maintaining an estimate of bandwidth available from the sender to the receiver; and

adjusting transmission based on the estimate.

2. The method according to claim 1, wherein transmission is in real time.

3. The method according to claim 1, wherein maintaining the estimate of bandwidth comprises monitoring of packet loss based on acknowledgments from the receiver.

4. The method according to claim 1, wherein, in maintaining the estimate of bandwidth, the sender maintains a count of packets outstanding.

5. The method according to claim 4, wherein, in maintaining the estimate of bandwidth, the sender maintains current an upper bound on how many packets are allowed to be outstanding.

6. The method according to claim 5, wherein the upper bound is as specified by the TCP congestion window.

7. The method according to claim 1, wherein, in maintaining the estimate of bandwidth, the sender maintains a count of bytes outstanding.

8. The method according to claim 7, wherein, in maintaining the estimate of bandwidth, the sender

maintains current an upper bound on how many bytes are allowed to be outstanding.

9. The method according to claim 8, wherein the upper bound is as specified by the TCP congestion window.

10. The method according to claim 1, further comprising retransmitting a packet which has been determined by the receiver as having been lost in transmission or received with an error.

11. The method according to claim 1, further comprising adapting bandwidth required by the data.

12. The method according to claim 1, further comprising discriminating between packets lost due to congestion in the network and packets received with at least one bit error.

13. A system for transmitting data from a sender to a receiver in a digital communications network, comprising:

means for maintaining an estimate of bandwidth available from the sender to the receiver; and

means for adjusting transmission based on the estimate.

14. The system according to claim 13, wherein transmission is in real time.

15. The system according to claim 13, wherein the means for maintaining the estimate of bandwidth comprises means for monitoring of packet loss based on acknowledgments from the receiver.

16. The system according to claim 13, wherein the means for maintaining the estimate of bandwidth comprises means for maintaining a count of packets outstanding.

17. The system according to claim 16, wherein the means for maintaining the estimate of bandwidth comprises means for maintaining current an upper bound on how many packets are allowed to be outstanding.

18. The system according to claim 17, wherein the upper bound is as specified by the TCP congestion window.

19. The system according to claim 13, wherein the means for maintaining the estimate of bandwidth comprises means for maintaining a count of bytes outstanding.

20. The system according to claim 19, wherein the means for maintaining the estimate of bandwidth comprises means for maintaining current an upper bound on how many bytes are allowed to be outstanding.

21. The system according to claim 20, wherein the upper bound is as specified by the TCP congestion window.

22. The system according to claim 13, further comprising means for retransmitting a packet which has been determined by the receiver as having been lost in transmission or received with an error.

23. The system according to claim 13, further comprising means for adapting bandwidth required by the data.

24. The system according to claim 13, further comprising means for discriminating between packets lost

due to congestion in the network and packets received with at least one bit error.

25. A system for transmitting data from a sender to a receiver in a digital communications network, comprising a processor which is instructed for:

maintaining an estimate of bandwidth available from the sender to the receiver; and

adjusting transmission based on the estimate.

26. The system according to claim 25, wherein transmission is in real time.

27. The system according to claim 25, wherein maintaining the estimate of bandwidth comprises monitoring of packet loss based on acknowledgments from the receiver.

28. The system according to claim 25, wherein, in maintaining the estimate of bandwidth, the sender maintains a count of packets outstanding.

29. The system according to claim 28, wherein, in maintaining the estimate of bandwidth, the sender maintains current an upper bound on how many packets are allowed to be outstanding.

30. The system according to claim 29, wherein the upper bound is as specified by the TCP congestion window.

31. The system according to claim 25, wherein, in maintaining the estimate of bandwidth, the sender maintains a count of bytes outstanding.

32. The system according to claim 31, wherein, in maintaining the estimate of bandwidth, the sender maintains current an upper bound on how many bytes are allowed to be outstanding.

33. The system according to claim 32, wherein the upper bound is as specified by the TCP congestion window.

34. The system according to claim 25, wherein the processor is instructed further for retransmitting a packet which has been determined by the receiver as having been lost in transmission or received with an error.

35. The system according to claim 25, wherein the processor is instructed further for adapting bandwidth required by the data.

36. The system according to claim 25, wherein the processor is instructed further for discriminating between packets lost due to congestion in the network and packets received with at least one bit error.

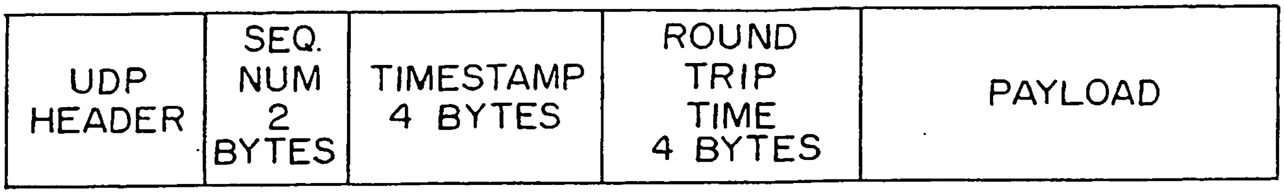


FIG. 1

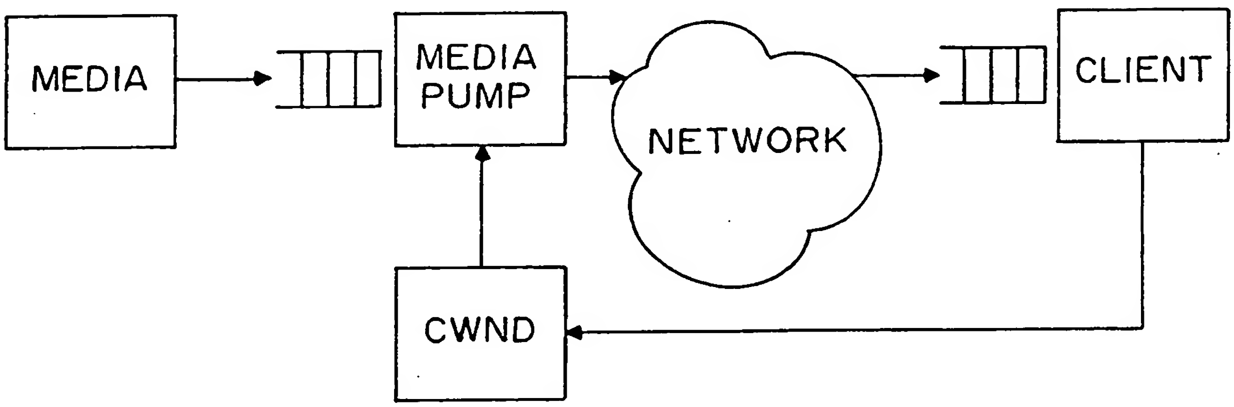


FIG. 3a

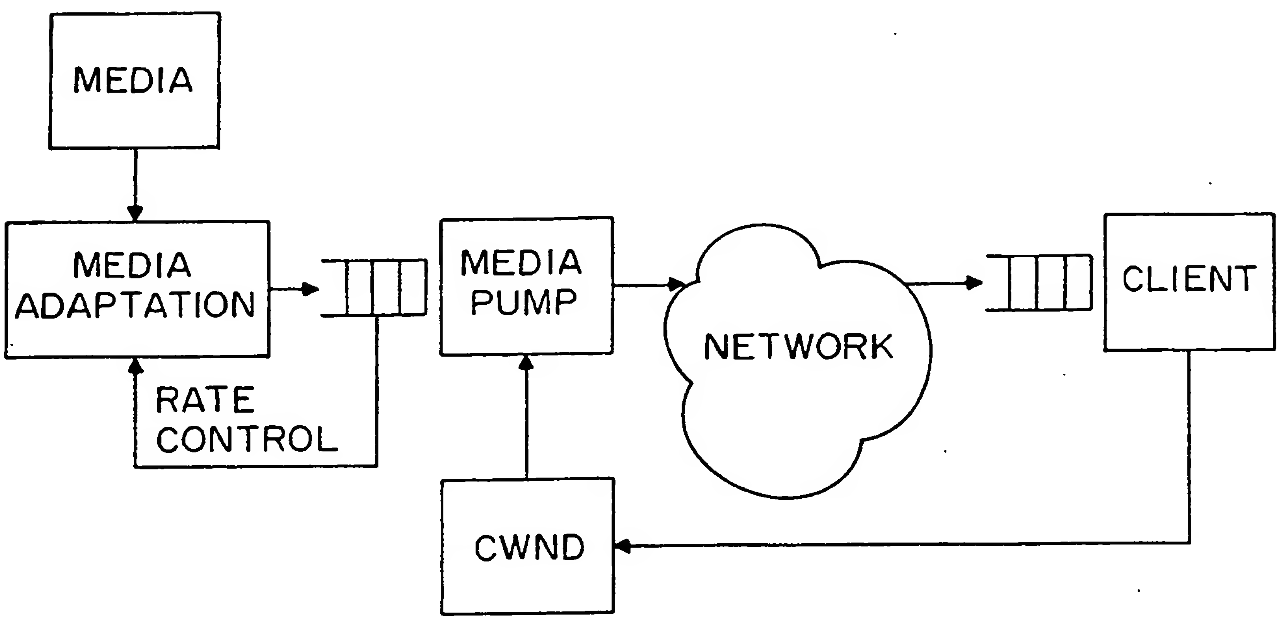


FIG. 3b

2/4

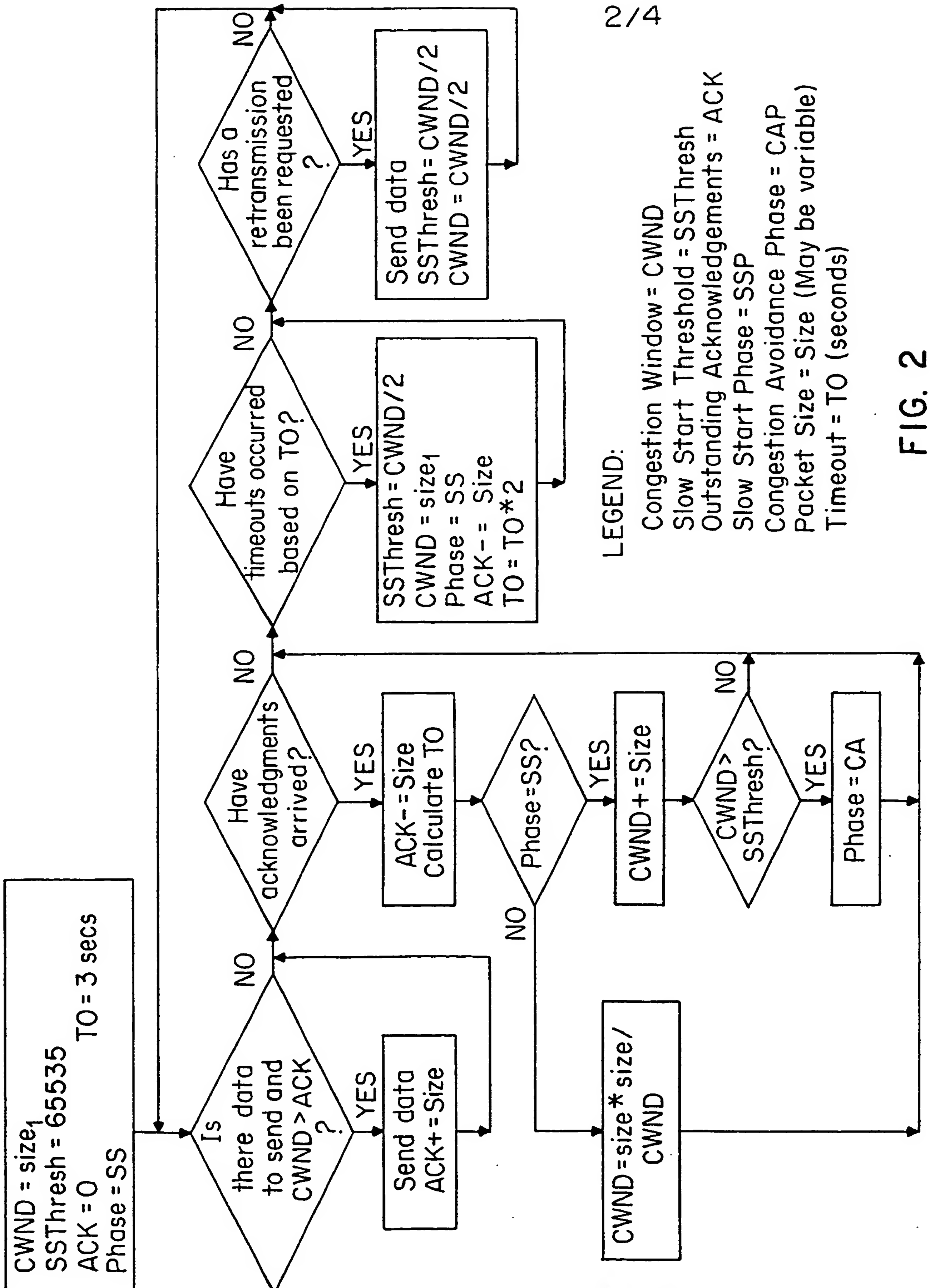


FIG. 2

SUBSTITUTE SHEET (RULE 26)

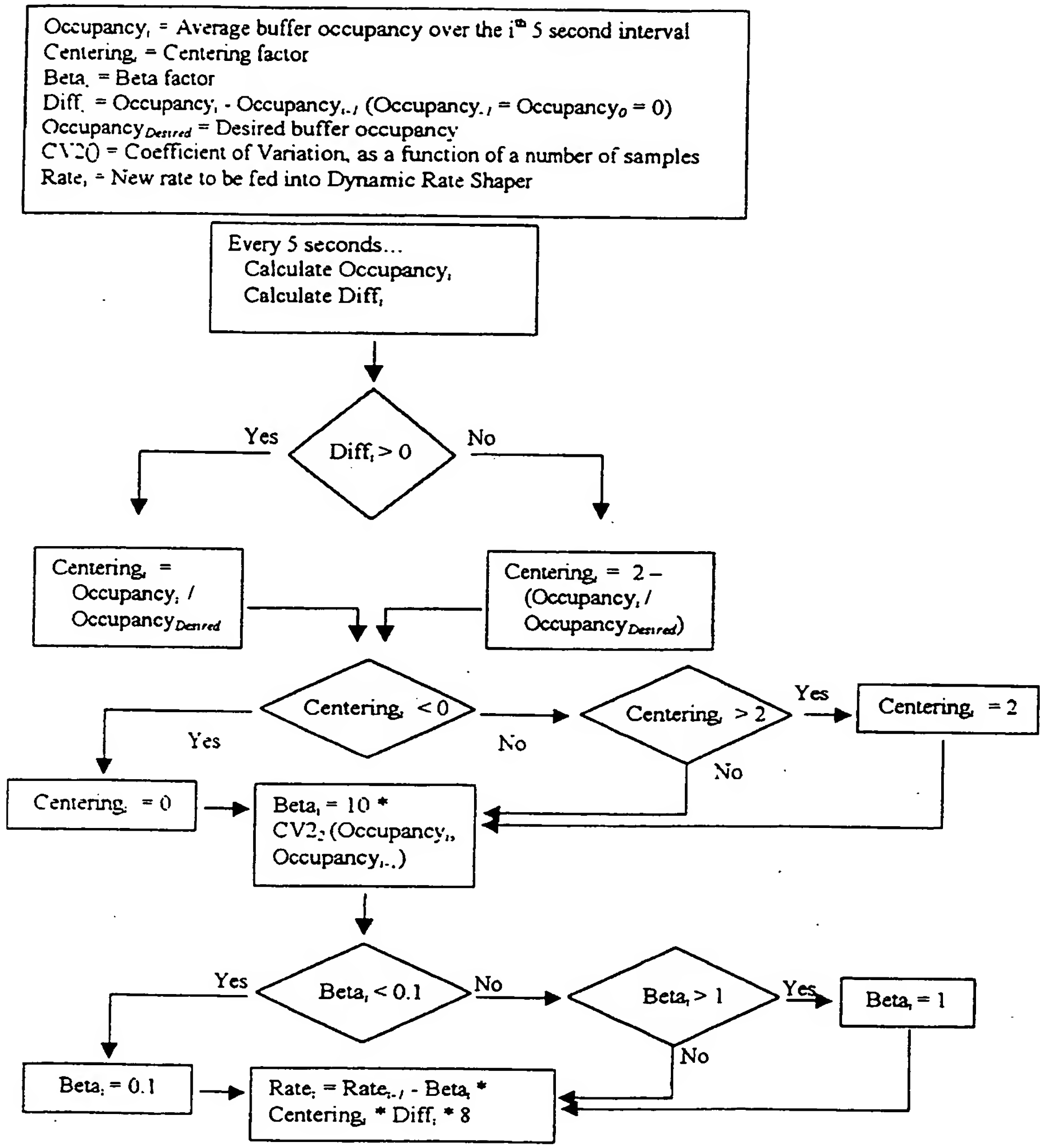

Occupancy, = Average buffer occupancy over the ith 5 second interval
Centering, = Centering factor
Beta, = Beta factor
Diff, = Occupancy, - Occupancy,-1 (Occupancy-1 = Occupancy0 = 0)
OccupancyDesired = Desired buffer occupancy
CV20 = Coefficient of Variation, as a function of a number of samples
Rate, = New rate to be fed into Dynamic Rate Shaper
Every 5 seconds...
Calculate Occupancy,
Calculate Diff,
Diff, > 0
Yes
No
Centering, = Occupancy, / OccupancyDesired
Centering, = 2 – (Occupancy, / OccupancyDesired)
Centering, < 0
Yes
No
Centering, > 2
Yes
No
Centering, = 2
Centering, = 0
Beta, = 10 * CV2, (Occupancy,, Occupancy,-1)
Beta, < 0.1
Yes
No
Beta, > 1
Yes
No
Beta, = 1
Beta, = 0.1
Rate, = Rate,-1 - Beta, * Centering, * Diff, * 8

Fig. 4

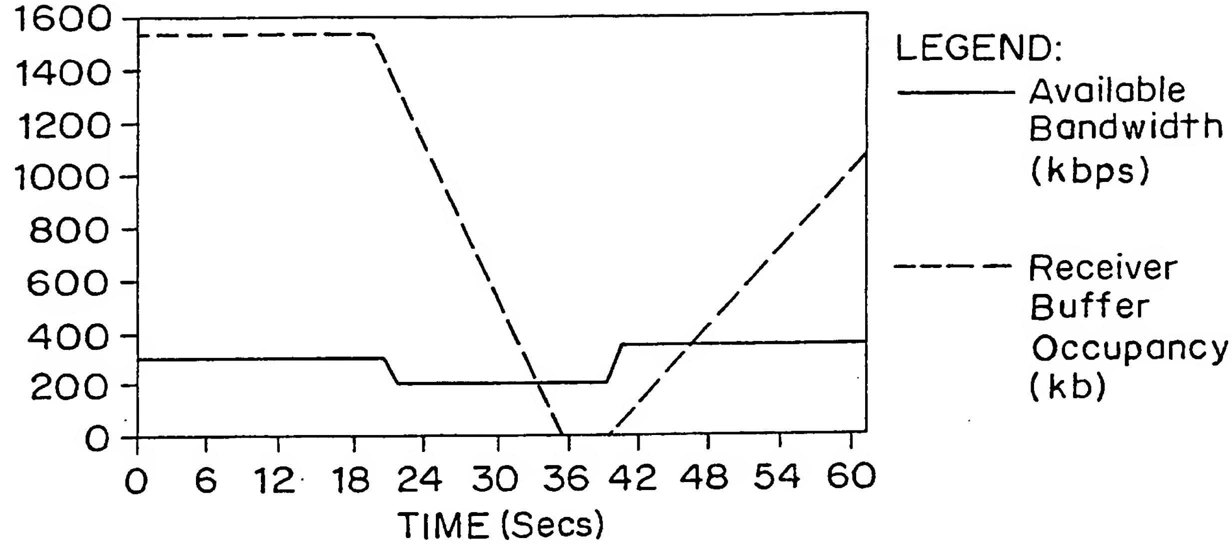


FIG. 5a

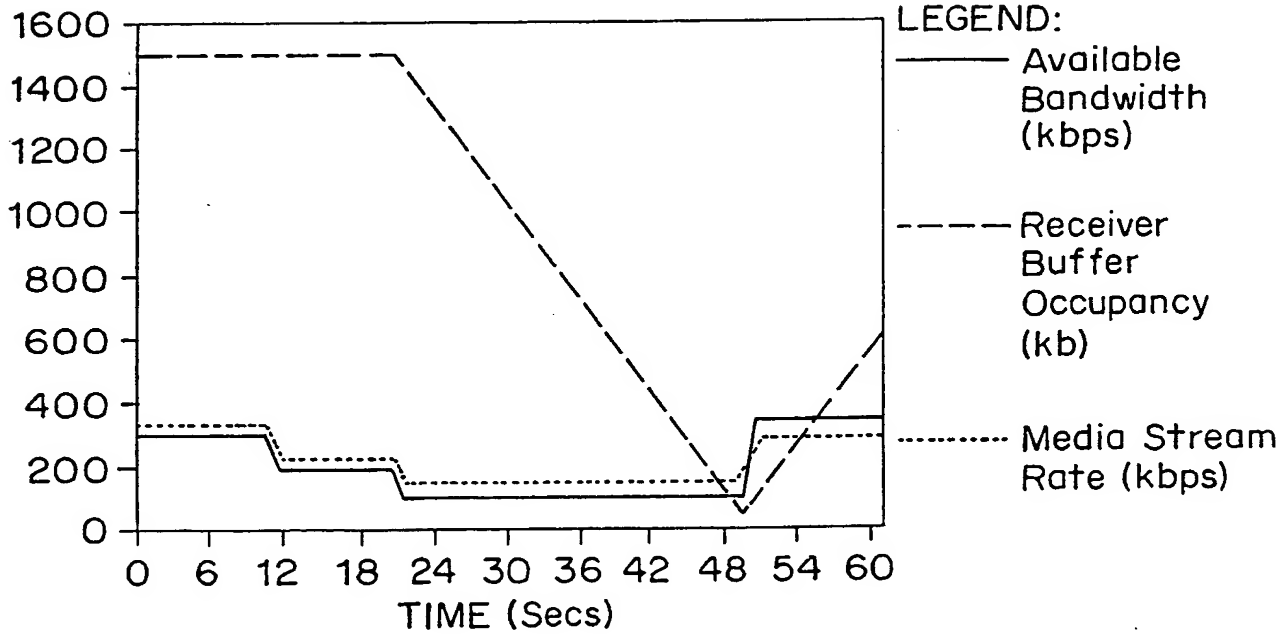


FIG. 5b

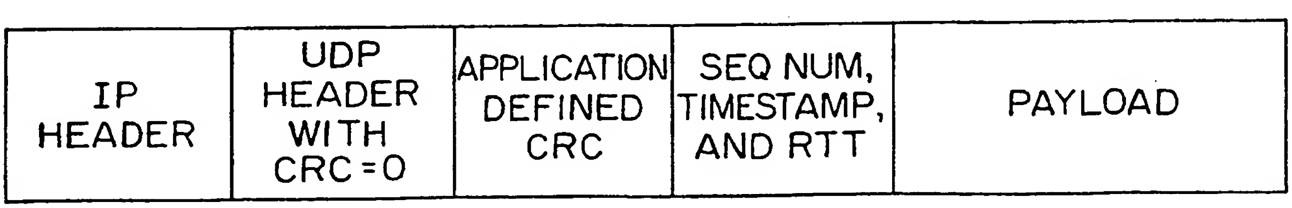


FIG. 6

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US97/19207

A. CLASSIFICATION OF SUBJECT MATTER

IPC(6) : H04J 3/16
US CL : 370/468
According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

U.S. : 370/252, 468, 477; 375/240

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	US 5,526,350 A (GITTINS et al) 11 June 1996, col. 7, lines 49-65.	1-36
X	US 5,627,970 A (KESHAV) 06 May 1997, col. 6, lines 9-31 & 46-65.	1-36
A	US 5,115,309 A (HANG) 19 May 1992, abstract.	1-36
A	US 5,490,252 A (MACERA et al) 06 February 1996, abstract.	1-36

☐ Further documents are listed in the continuation of Box C. ☐ See patent family annex.

* Special categories of cited documents:

"A" document defining the general state of the art which is not considered to be of particular relevance

"B" earlier document published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

"&" document member of the same patent family

Date of the actual completion of the international search

08 JANUARY 1998

Date of mailing of the international search report

17 FEB 1998

Name and mailing address of the ISA/US
Commissioner of Patents and Trademarks
Box PCT
Washington, D.C. 20231

Facsimile No. (703) 305-3230

Authorized officer
RICKY QUOC NGO

Telephone No. 703-305-4798